JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

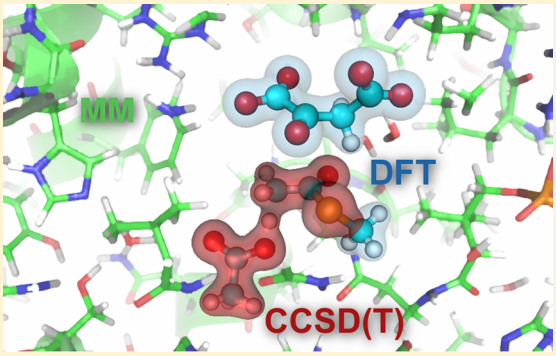
A Projector-Embedding Approach for Multiscale Coupled-Cluster Calculations Applied to Citrate Synthase

Simon J. Bennie,*[,†] Marc W. van der Kamp,[†,§] Robert C. R. Pennifold,[†] Martina Stella,[†,#] Frederick R. Manby,*[,†] and Adrian J. Mulholland*[,†]

[†]Center for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

[§]School of Biochemistry, Biomedical Sciences Building, University Walk, University of Bristol, Bristol BS8 1TD, U.K.

ABSTRACT: Projector-based embedding has recently emerged as a robust multiscale method for the calculation of various electronic molecular properties. We present the coupling of projector embedding with quantum mechanics/molecular mechanics modeling and apply it for the first time to an enzyme-catalyzed reaction. Using projector-based embedding, we combine coupled-cluster theory, density-functional theory (DFT), and molecular mechanics to compute energies for the proton abstraction from acetyl-coenzyme A by citrate synthase. By embedding correlated *ab initio* methods in DFT we eliminate functional sensitivity and obtain high-accuracy profiles in a procedure that is straightforward to apply.



INTRODUCTION

Electronic structure calculations are becoming essential for identifying and testing reaction mechanisms in enzymes[1,2] and are finding ever wider application in biology and biochemistry. To reach accurate chemical results in biological systems a common approach is to combine a quantum mechanics (QM) treatment of the reacting center of an enzyme combined with a molecular mechanics (MM) approach for the enzyme environment (QM/MM).[1,3,4] Such calculations allow for the examination of short-lived species, such as transition states and reaction intermediates, that are often not directly accessible through experimental methods.[5,6]

It is known that for some enzyme reaction mechanisms Hartree−Fock (HF) theory can give barriers up to two times higher than experiment.[4,7] Second-order Møller−Plesset perturbation theory (MP2) is more accurate, but tends to underestimate barriers.[8,9] Density functional theory (DFT) is an attractive alternative, but its accuracy can be hard to predict, and often it is useful to validate models against coupled-cluster calculations.[10,11] Reliable calculations require high-level methods: "chemically accurate" (within 1 kcal mol^{-1} of experiment) predictions of reaction barriers require coupled-cluster theory. Such accurate QM calculations are possible using local coupled cluster methods (e.g., LCCSD(T0)).[4,7,12] An alternative is to use (nonlocal) spin-component-scaled MP2 (SCS-MP2)[13] in QM/MM calculations,[14,15] as this method correlates reasonably well with LCCSD(T0) results for reaction barriers, often compensating for the underestimation of MP2.[7,12]

DFT is successfully used as an efficient QM method in many research communities, but picking the "right" approximate exchange-correlation functional can be problematic, and many popular approximations lead to considerable errors in calculated reaction barriers.[7−9] Despite ongoing improvements to density functional approximations,[16−21] the choice of functional can strongly affect the calculated result.[22−24] While in some cases DFT errors may be limited to underestimation of energy barriers for reactions, in others, barriers may be overestimated. In some cases, DFT with many standard functionals leads to qualitatively incorrect conclusions, for example by predicting the wrong mechanism for some enzyme-catalyzed reactions.[9,25] DFT can also give poor prediction of spectroscopic properties in transition-metal complexes, such as spin-state and zero-field splittings, and g-tensors.[26−28] *Ab initio* methods have the potential to be more accurate.[28]

In contrast to *ab initio* wave function methods, no scheme for systematic improvement of DFT has been found. Most functionals lack an inherent description of van der Waals interactions arising from dispersion (although various corrections are available to overcome this; the most commonly used is the empirically corrected B3LYP-D and related variants[29−31]). Inclusion of dispersion effects has been shown to improve DFT results for modeling enzyme-catalyzed reactions.[15,32−34] Also, the self-interaction error in DFT can result in unphysical delocalization of orbitals.[35−37] *Ab initio* theories remove ambiguity in the choice of functional and improves the predictive power of QM/MM reaction calculations. Ideally, approaches to remove such ambiguity should fit in with standard workflows and have a low overhead for adoption.

For large systems such as proteins even the relatively low scaling of DFT[38,39] or local correlation methods becomes a problem, although linear scaling implementations are avail-

Received: March 21, 2016
Published: May 9, 2016

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

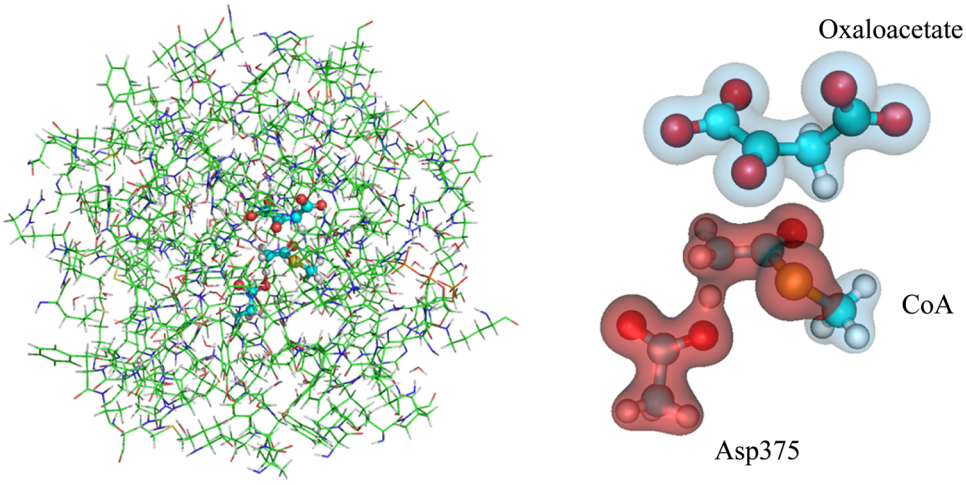


Figure 1. Transition state structure of the first step in the reaction catalyzed by citrate synthase: proton abstraction from acetyl-coenzyme A by Asp375.[7] Left: The model of chicken citrate synthase used in this study (truncated to a 15 Å sphere, and previously optimized at the B3LYP/6-31+G(d,p)//CHARMM27 level). The protein environment (treated with MM) is depicted with bonds as lines (carbon atoms in green) and encompassing subsystems A, B, and C. Right: The QM subsystem with the densities of subsystem A are highlighted in red and of subsystem B in blue.

able.[40] Two routes around this are possible: the system can be truncated into a model system with the assumption that the removed moieties do not affect the electronic structure, or one can utilize a multiscale approach that is capable of combining speed and accuracy with the aim of capturing the majority of long-range effects. The most common multilevel approach is the QM/MM approach pioneered by Karplus, Warshel, and Levitt.[1,3,41,42] Many applications of this approach in biomolecular systems use DFT in a subsystem of chemical significance for the highest accuracy energy calculations, while the rest of the system is treated with a low-cost MM empirical "force field". The coupling between the regimes is typically achieved by polarizing the QM one-electron Hamiltonian by the MM point charges and using (MM) Lennard-Jones terms for van der Waals interactions.[1,42] This simple QM/MM approach can give good results for biomolecular interactions[43] and reactions.[4] It can be necessary to optimize Lennard-Jones parameters for QM/MM interactions[44] and the consistency and compatibility of QM and MM methods should be considered.[45] In some cases artifacts can arise in QM/MM, such as overpolarization due to the point charge approximation, although newer force fields are actively being developed.[46−48] While it is common to check that answers are converged with respect to the size of the quantum subsystem,[49] it does not guarantee that the quantum method has captured the underlying chemistry.[15,25,50] We aim to address this by highlighting a method to calculate high-level quantum chemical properties, which also provides an accurate description of the environment through quantum polarization that is obtained from standard DFT.

Wave function (*ab initio*) methods allow for precise conclusions to be drawn about mechanistic pathways and spectroscopic assignments. Another attractive feature is that wave function methods allow for an easy interpretation of the sources of error due to their systematic improvability. Wave function methods such as the complete active space self-consistent field (CASSCF) method are able to tackle strongly correlated problems such as when a near degeneracy of the HOMO−LUMO (highest occupied molecular orbital−lowest unoccupied molecular orbital) occurs (i.e., bond stretching and transition metal complexes), which are usually out of the reach of DFT.[28] The cost of utilizing accurate wave function methods such as the gold standard CCSD(T) is that they scale poorly with system size (N^7). While approximations exist that take advantage of the local nature of correlation to significantly improve scaling,[51−53] it is attractive to have a method that can return accurate properties for all post-HF methods from MP2 all the way up to full configuration−interaction without strongly scaling cost with the size of the surrounding environment.

Several methods exist that are forms of quantum embedding: a simple example is the frozen core approximation for wave function correlation.[54] Embedding methods can be constructed using the one-particle density matrix,[55,56] Green's functions,[57] or, in a DFT formalism, the electronic density.[58−63] An alternative is to use region-based local coupled-cluster methods.[64]

This work focuses on the use of a conceptually simple and computationally cheap embedding technique that uses projection operators to exactly embed high-accuracy wave function methods inside DFT potentials, here called WF-in-DFT; further we incorporate this embedded treatment of the active site into an MM treatment of the wider protein environment. This multilevel approach enables us to utilize the most accurate wave function methods for chemically critical regions (subsystem A; see Figure 1), more approximate DFT methods (subsystem B) for chemical moieties that are not directly involved in the reaction but that can affect the electronic structure of subsystem A, and MM for the longer range effects from the protein environment (subsystem C).

For this study, we examine a well-characterized enzyme-catalyzed reaction, which has become a testbed for computational modeling,[7,9,65−67] namely the deprotonation of the acetyl coenzyme A (acetyl-CoA) in citrate (*Si*-)synthase. Overall this reaction forms citric acid from acetyl-CoA and oxaloacetate, the first step of the citric acid cycle. Citrate synthase is crucial for most forms of life, and it has been studied extensively both experimentally[68−71] and computationally.[7,9,65−67] The first step in the reaction mechanism of citrate synthase involves proton abstraction from the α-carbon of acetyl-CoA by an aspartate

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

residue (Asp375, in the numbering for the pig enzyme). The enolate intermediate that is thus formed[7,67] then performs a nucleophilic attack on the carbonyl carbon of oxaloacetate, forming a stable citryl-CoA intermediate. Finally, hydrolysis of the thioester bond produces citrate and coenzyme A. For efficient catalysis of the overall reaction, stabilization of the enolate intermediate by the enzyme active site is crucial.[7,9] Here, we examine the performance of WF-in-DFT calculations coupled to molecular mechanics calculations, leading to unprecedented canonical CCSD(T)-level QM/MM calculations on an enzyme catalyzed reaction.

THEORY

To embed wave function methods in DFT, we utilize a projector-based embedding scheme that has the property that the DFT-in-DFT (e.g., PBE-in-PBE) energy is the same as the full DFT calculation.[72] This is achieved without having to resort to numerically challenging, iterative procedures such as optimized effective potentials.[59−63] A further advantage of this method is that it only requires modification of the core Hamiltonian, so practically any wave function method can be used in the high-accuracy subsystem.

The total electronic density (ρ^{T}) can be directly partitioned into subsystem contributions as

$$\rho^{\mathrm{T}} = \rho^{\mathrm{A}} + \rho^{\mathrm{B}} \tag{1}$$

The DFT energy decomposes into contributions from the two subsystems and a nonadditive component that describes the interaction between them:

$$E[\rho^{\mathrm{T}}] = E[\rho^{\mathrm{A}}] + E[\rho^{\mathrm{B}}] + \delta E[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] \tag{2}$$

The nonadditive component is made up of the nonadditive Coulomb, kinetic, and exchange-correlation energies; of these terms, the kinetic energy is the most challenging contribution.[72] Although methods exist to calculate this term, they are either expensive[59−63] or are not yet accurate enough to be used for partitions that cut covalent bonds.[73−77]

In projector-based embedding[72,78−82] the requirement to calculate the kinetic energy contribution is avoided by constructing the subsystem densities from orthogonal subsets of orbitals: thus, the nonadditive kinetic energy interaction is exactly zero. In our implementation, orthogonality is achieved by applying a projection operator that elevates all subsystem B orbitals to high energy, making them unavailable to the electrons of subsystem A. The subsystem-A core Hamiltonian is given by

$$\mathbf{h}^{\mathrm{A\,in\,B\,in\,C}} = \mathbf{h} + \mathbf{J}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] - \mathbf{J}[\rho^{\mathrm{A}}] + \mathbf{v}_{\mathrm{xc}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] - \mathbf{v}_{\mathrm{xc}}[\rho^{\mathrm{A}}] + \mu\mathbf{P}^{\mathrm{B}} + \mathbf{h}^{\mathrm{C}}_{\mathrm{QM/MM}} \tag{4}$$

where $\mathbf{h}$ is the original core Hamiltonian; $\mathbf{J}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] - \mathbf{J}[\rho^{\mathrm{A}}]$ describes the Coulomb interaction with electrons of subsystem B; $\mathbf{v}_{\mathrm{xc}}[\rho^{\mathrm{A}}, \rho^{\mathrm{B}}] - \mathbf{v}_{\mathrm{xc}}[\rho^{\mathrm{A}}]$ similarly describes the exchange correlation interaction; $\mathbf{P}^{\mathrm{B}}$ is the projector onto orbitals in B; μ is a large, positive parameter; and $\mathbf{h}^{\mathrm{C}}_{\mathrm{QM/MM}}$ is the usual Hamiltonian describing electrostatic interaction with the MM charges in subsystem C.

The total energy for a WF-in-DFT calculation is given by

$$E[\psi^{\mathrm{A}}; \gamma^{\mathrm{A}}, \gamma^{\mathrm{B}}] = \langle\psi^{\mathrm{A}}|\mathbf{h}^{\mathrm{A\,in\,B\,in\,C}}|\psi^{\mathrm{A}}\rangle - \mathrm{tr}\,\gamma^{\mathrm{A}}(\mathbf{v}_{\mathrm{xc}}[\gamma^{\mathrm{A}}, \gamma^{\mathrm{B}}] - \mathbf{v}_{\mathrm{xc}}[\gamma^{\mathrm{A}}]) + E_{\mathrm{xc}}[\gamma^{\mathrm{A}}, \gamma^{\mathrm{B}}] - E_{\mathrm{xc}}[\gamma^{\mathrm{A}}] + E[0, \gamma^{\mathrm{B}}] + E^{\mathrm{coupling}}_{\mathrm{QM/MM}} + E^{\mathrm{C}}_{\mathrm{MM}} \tag{5}$$

where $\mathbf{h}^{\mathrm{A\,in\,B\,in\,C}}$ is the embedding Hamiltonian in eq 4 used by the high-level wave function method; the central terms describe a correction for exchange-correlation double-counting;[72] and $E^{\mathrm{coupling}}_{\mathrm{QM/MM}}$ and $E^{\mathrm{C}}_{\mathrm{MM}}$ are the QM/MM contributions from subsystem C.

The canonical Kohn−Sham orbitals tend to be delocalized over the system, so to define subsystems the orbitals are first localized and then automatically assigned to atomic centers by population analysis. Previous work has shown that the intrinsic bond orbitals (IBOs) of Knizia[83] provide reliable and chemically meaningful localized orbitals.[78]

To substantially reduce the size of the virtual space in the WF calculation on subsystem A, we also use basis truncation to reduce the number of two-electron integrals (and wave function amplitudes) in a systematic and controllable fashion.[78] Asymptotically, this leads to the computational cost of the high-level calculation in subsystem A becoming independent of the size of the environment. Two truncation methods are available:[78,81] we use the form proposed by Bennie et al.[78] A single parameter based on the net Mulliken population is used to screen atomic orbitals; for various chemical systems we were able to achieve submillihartree truncation errors with the threshold set to 10^{-4}.

The procedure for CCSD(T)-in-DFT/MM requires only the definition of atomic centers in the active (A) subsystem as an extra step in the input. Overall, the procedure takes the following form:

1. Take a molecular structure (e.g., previously optimized with QM/MM).
2. Perform single point DFT on the QM subsystem (A+B).
3. Localize the orbitals.
4. Select the atoms to be in subsystem A for the high level (e.g., CCSD(T)) calculation.
5. Perform high-level (e.g., CCSD(T)) calculations on subsystem A, using the embedding Hamiltonian given in eq 4, which results in the embedded WF-in-DFT/MM energy.

METHODS

To study the proton abstraction of acetyl-CoA by Asp375 in citrate synthase with projector embedded QM/MM, we used previously optimized QM/MM geometries (at the B3LYP/6-31+G(d)//CHARMM27 level).[7] These were obtained by employing a reaction coordinate defined as RC = $d(\mathrm{C_{Acetyl\text{-}CoA}}\text{−H}) - d(\mathrm{O_{Asp375}}\text{−H})$. This reaction coordinate has been shown to represent the reaction pathway accurately (e.g., by comparison to the nudged elastic band method).[7]

We used geometries that were obtained between RC = −1.4 Å and RC = 1.4 Å with a step size of 0.1 Å. For the majority of the calculations, the QM region consisted of the methylthioester part of acetyl-CoA, the side chain of Asp375, and oxaloacetate (OAA) (see the right-hand side in Figure 1). We also performed calculations with larger QM regions (regions 2−4 in Figure 4). Single-point QM/MM and WF-in-DFT/MM calculations were performed using the Molpro 2015.1 software

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

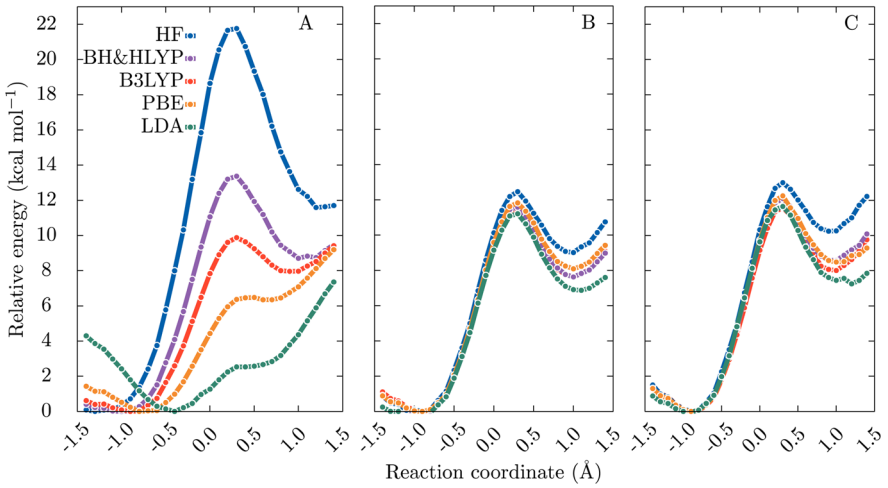


Figure 2. QM/MM potential energy profiles of the proton transfer from acetyl-CoA to Asp375 in citrate synthase: comparison of different methods for embedding. Left (A): Results with various traditional DFT methods and HF in subsystems A and B (no correlated wave function method employed in subsystem A). Center (B): The CCSD(T) embedding barrier with the default basis truncation threshold (10^{-4}) in various DFT functionals and HF. Right (C): The embedding barrier with an energetically looser (computationally cheaper) truncation of the basis for various functionals; truncation level 10^{-3}. The basis set was aug-cc-pVDZ in all cases.

package.[84,85] The MM energy and QM/MM van der Waals contributions were taken from the previous B3LYP/6-31+G-(d)//CHARMM27 calculations.[7]

All the calculations in this work used the aug-cc-pVDZ or aug-cc-pVTZ basis sets[86−88] for all atoms in subsystems A and B. QM/MM calculations were performed using the LDA,[89,90] PBE,[91] B3LYP,[92] and BH&HLYP[93] functionals, as well as HF theory, for the QM subsystem. WF-in-DFT/MM calculations used each of these mean-field methods for subsystem B and CCSD(T),[94,95] MP2,[96] SCS-MP2,[13] and CCSD levels on subsystem A.

The smallest QM region size we consider here is the equivalent of "region 2" in the previous high-level QM/MM work;[7] we designate it as our "region 1". This selection was made in order to have a sufficient region to show the effect of the choice of the functional: here the oxaloacetate molecule is included in subsystem B because QM treatment of oxaloacetate has been shown to influence the reaction profile significantly.[7] Region 1 is then subdivided into subsystems A and B, representing the CCSD(T) and DFT subsystems.

When choosing the size of subsystem A in embedding calculations it is important to ensure that enough electrons are included to obtain an accurate description of WF correlation in the reacting moiety. In the present citrate synthase case we use a selection that we have found to be widely applicable: subsystem A consisted of the three atoms directly involved in the reaction, plus all atoms up to two bonds away from these. This region is shown in Figure 1 in red, and contains 72 electrons, with the sulfur valency satisfied in the embedded calculation by including the covalent bond with the methyl group. The 1s, 2s, and 2p electrons on the sulfur atom were treated as core electrons and were not correlated in the post-HF methods. All other electrons, including the 1s electrons on carbon, nitrogen, and oxygen atoms, were correlated. The basis-set truncation scheme uses a threshold to determine how many basis functions to discard. With the threshold set to 0 the entire basis is retained, and as the threshold in increased, more subsystem B functions are discarded. The threshold values used for this study were 10^{-4} and 10^{-3}. All timing comparisons were conducted in serial on a dedicated workstation with 32GB of RAM. Density plots were generated by the Molekel 5.4 program[97] with an isovalue of 0.05. Enzyme geometries were visualized using the PyMOL software package.[98]

RESULTS AND DISCUSSION

Previous LCCSD(T0) QM/MM calculations indicate that the potential energy barrier for the proton abstraction in this structure is 11 kcal mol^{-1} at the complete basis set limit.[7] The transition state was found to be between 0.2 and 0.3 Å along the reaction coordinate. Although the oxaloacetate is merely a spectator in the enolate formation, it has a strong effect due to its proximity, and its inclusion in the QM region was found to be essential (treating it at only the MM point charge level leads to an increase in barrier height of ∼3 kcal mol^{-1}).[66]

We first explore the reaction barrier and reaction energies using various traditional methods. Our selection is broad and includes the rather primitive local density approximation (LDA) functional, a general gradient approximation functional (PBE), two hybrid functionals (B3LYP and BH&HLYP), as well as HF theory. The resulting energy profiles vary significantly (Figure 2A) in line with previous work.[4,7] For example, a barrier of 20 kcal mol^{-1} was found when using HF while LDA gives a small barrier of 2.5 kcal mol^{-1}. Pure DFT calculations underestimate the barrier and do not strongly indicate a minimum for the enolate intermediate. BH&HLYP is 4 kcal mol^{-1} closer to the HF answer compared to B3LYP, presumably due to the higher proportion of exact exchange (50% vs 20%).

Using projector embedding to perform CCSD(T) in subsystem A (Figure 2B) largely removes the variation due to choice of functional in subsystem B. Across all methods the energy profiles are now quantitatively similar, particularly for the (approximate) activation energy, which has a maximum variance of 1 kcal mol^{-1} between B3LYP and HF. Variation for the reaction energy of the enolate intermediate with respect to the reactant state is somewhat larger. This variance is due to HF and LDA, which may be producing poorer oxaloacetate densities, and thus failing to model the stabilization of the CoA by the oxaloacetate ketone. The GGA and hybrid functionals (PBE, B3LYP and BH&HLYP) are in much closer agreement

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

for the reaction energy, with a variation of under 0.4 kcal mol^{-1}, compared to 6 kcal mol^{-1} without embedding. Notably, the locations of the reactant state minima, the transition state maxima, and the enolate intermediate minima are very independent of functional in the embedding calculations (reaction coordinate values of −1.0, 0.3, and 1.0 Å, respectively). This contrasts to the results without embedding, with which widespread values are found for the reactant minimum, between −1.3 (HF) and −0.4 Å (LDA), and for the product minimum, which ranges between 0.6 and 1.1 Å. This again highlights the consistency of projector-based WF-in-DFT embedding compared to traditional DFT.

Previously, LCCSD(T0)/aug-cc-pVDZ//MM energy calculations with the same geometries (with LCCSD(T0) employed for subsystems A and B) indicated an activation barrier of 13.2 kcal mol^{-1} and a reaction energy of 8.4 kcal mol^{-1}. Here, CCSD(T)-in-DFT calculations with all functionals for the environment were found to be within 2 kcal mol^{-1} of LCCSD(T0) for the activation barrier, and the calculations using PBE, BH&HLYP, and B3LYP were within 1 kcal mol^{-1} for the reaction energy. The CCSD(T)-in-DFT methods resulted in a barrier of between 11.4 and 12.1 kcal mol^{-1}. This includes CCSD(T)-in-LDA, which in sharp contrast to the unembedded LDA result has both the correct qualitative shape for the transition state as well as barrier height. The HF method resulted in a barrier 0.3−1.0 kcal mol^{-1} higher than obtained with the density functionals. A contributing factor to this difference is probably that the underlying HF method lacks correlation and therefore provides a poorer embedding environment; this is also reflected in the HF barrier being a factor of around two too high in Figure 2A. Using B3LYP and BH&HLYP as the subsystem B method resulted in reaction profiles that differed by a maximum of 0.1 kcal mol^{-1}, showing that the density from a reasonable functional results in almost identical barriers, despite the intrinsic difference between these two functionals being up to 4 kcal mol^{-1}.

The basis truncation default of 10^{-4} reduced the number of system basis functions from 517 contracted functions to 412, which corresponds to 62% of the environment (Figure 2B). The looser threshold (10^{-3}) retains 352 functions, corresponding to 40% of the environment (Figure 2C). We find little effect on the barrier height despite a reduction in computational cost by a factor of 2 (averaged over the profile). It should be noted that the enolate minimum energy differs by, at most, 1.2 kcal mol^{-1} for HF, but for PBE and B3LYP is only 0.2 kcal mol^{-1} higher compared to the tighter truncation threshold of 10^{-4}. Calculations with a looser threshold (such as 10^{-3}) may be useful to explore the performance of WF-in-DFT/MM calculations, but it is advisible to use an energetically tighter threshold (such as the default 10^{-4}) for final results.

Projector embedding only requires a modified core Hamiltonian, so it is possible to use essentially any electronic structure method in subsystem A. We can thus compare how various approaches to include electron correlation in subsystem A affect the reaction energy profile (Figure 3; PBE is used in subsystem B). As expected, the absence of electron correlation (HF) leads to the largest barrier of 20 kcal mol^{-1}. It should be noted that HF embedded in PBE (Figure 3) leads to a different profile than unembedded HF (Figure 2A), this is because HF is "restrained" by the PBE environment, leading to a decrease of the barrier height of 1.6 kcal mol^{-1}. When taking CCSD(T) as the reference, MP2 underestimates the activation and reaction energies by 1.7 kcal mol^{-1} and 1.0 kcal mol^{-1}, respectively.

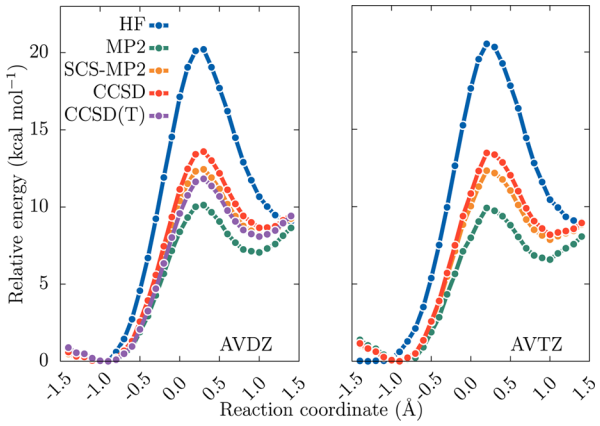


Figure 3. WF-in-DFT/MM potential energy profiles of the proton transfer from acetyl-CoA to Asp375 in citrate synthase: Effect of different WF methods. In all cases, the WF method is embedded in a PBE environment. Left: aug-cc-pVDZ basis on all atoms. Right: aug-cc-pVTZ basis on all atoms.

Spin-component-scaled MP2 (SCS-MP2) is more accurate with activation and reaction energies 0.6 kcal mol^{-1} and 0.1 kcal mol^{-1} higher than CCSD(T), respectively. The relative performance of MP2 and SCS-MP2 compared to CCSD(T) is consistent with what was found previously for the same reaction[7] and the two other enzyme reactions studied at this level,[4,12] with (L)MP2 underestimating activation barriers and SCS-(L)MP2 showing activation barriers close to the LCCSD(T0) result. Increasing the basis size from aug-cc-pVDZ to aug-cc-pVTZ was found to decrease the barrier height by 0.1−0.3 kcal mol^{-1} and to move the transition state to RC = 0.2 Å for all methods other than HF.

Previous work by Van der Kamp et al. showed the SCS-MP2 results to be insensitive to the basis size, while the LCCSD(T0) results were found to be more strongly affected by the size of basis, changing by up to 1.5 kcal mol^{-1} with aug-cc-pVTZ. Our results here agree with the canonical SCS-MP2 of that work, in that the basis size does not have a large effect on barrier height. It is possible that the effect of increasing basis set size found previously for LCCSD(T0) calculations may result from a coupling between domain completeness and basis-set size. It has been shown that the basis dependence of LCCSD(T0)/aug-cc-pVTZ calculations of activation barriers for the hydroxylation reaction catalyzed by *p*-hydroxybenzoate hydroxylase can be circumvented by performing LCCSD/aug-cc-pVTZ with a coupled cluster triples correction in the smaller (aug)-cc-pVDZ basis but with the same domains as in the larger basis.[12] It may thus be "good practice" to use a large basis with older domain-based local correlation methods to avoid an artificially high barrier due to basis-set incompleteness effects, or more advisably newer local methods with F12 explicit correlation should be used.[99−102]

Truncated projector embedding has the attractive property that the cost of the coupled-cluster calculation is not (formally) linked to the size of subsystem B, provided that the active subsystem A density does not become more delocalized over the system as more fragments are added to subsystem B. The decoupling occurs because the number of atomic functions needed to describe the molecular orbitals in the active subsystem becomes constant as atomic functions on increasingly distant atoms are added. Figure 4 shows that for the smallest region (encompassing the methylthioester part of

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

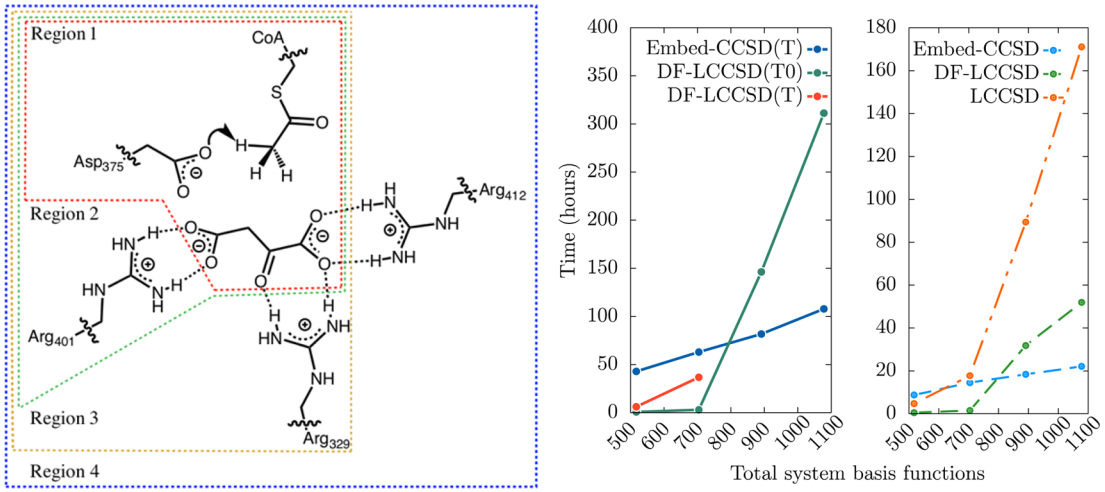


Figure 4. Timing benchmarks for increasing number of basis functions. Left: The molecular moieties encompassed by the increased region size going from the smallest (region 1) to the largest (region 4). Centre: Timings for single point LCCSD(T0)/aug-cc-pVDZ (with density fitting) and truncated PBE embedded-CCSD(T)/aug-cc-pVDZ for various region sizes. Right: Integral direct CCSD, density fitted LCCSD, and full 4-index integral LCCSD under the same benchmarking conditions. Benchmarking was conducted on a dedicated workstation with 32GB of RAM and the execution was performed in serial on an Intel Core i7−4790K CPU clocked at 4 GHz.

acetyl-CoA, the side chain of Asp375, and oxaloacetate), the DF-LCCSD(T0), with a noniterative perturbative treatment of triple excitations,[103] has a lower cost than embedded-coupled cluster with full perturbative treatment of triple excitations (T). Once the region grows to encompass over ∼800 basis functions, the lack of coupling to the environment size means that embedding remains moderate in cost. When using full triples, the local coupled cluster (DF-LCCSD(T)) calculation has a much higher cost.

Despite the fact that embedding provides a reasonable degree of decoupling between the active subsystem and the DFT environment, some atomic orbitals from neighboring atoms are needed, both for flexibility in the active subsystem wave function, and for proper representation of the projection operator. This effect can be minimized by using larger basis sets in the active subsystem[78] or by neglecting diffuse functions in the environment.

The right-hand side plot of Figure 4 shows the effect when neglecting the triples and doing the embedded CCSD calculation with direct integral evaluation. Changing from the density fitted[104−106] LCCSD (dashed green) to the full 4-index integrals (dashed orange) causes a ∼10-fold increase in cost for the local correlation method. These results indicate that a similar speedup would be possible if one was to embed a density-fitted wave function, although the best approach to obtain optimal scaling would be to use one of the CCSD(T)-F12[107,108] methods that are already available in Molpro with density fitting. When the truncated basis has large numbers of virtual functions retained we would expect an embedded density-fitting regime to offer some advantage. We intend to explore this in future work along with embedding some state-of-the-art local methods.

CONCLUSIONS

We have shown that projector-based WF-in-DFT embedding, coupled with basis-set truncation applied to the QM component of the calculation, can be applied to large biomolecular complexes, with little additional setup required relative to conventional DFT-based QM/MM calculations. The variability of results with respect to choice of approximate exchange-correlation functional for the reaction profile has been shown to be largely eliminated by embedding CCSD(T)-in-DFT, even when CCSD(T) is only applied to a small number of reacting atoms.

We also found that the basis truncation employed for the embedding does not affect the barrier height and our results are in line with previous LCCSD(T0) calculations. An important aspect of this work is that for larger quantum subregion sizes, projector embedding has reasonable scaling, meaning that it is possible to have large amounts of electronic polarization from DFT without greatly increasing the WF calculation cost. Second we are able to use any correlation method in the active subsystem, thus opening up the potential for using powerful new multireference methods in biomolecular simulations that call for the treatment of static correlation. CCSD(T)-in-DFT/MM provides an alternative way to calculate coupled-cluster level QM/MM reaction pathways that uses established knowledge about DFT and corrects for its limitations. A DFT treatment of the environment surrounding the CCSD(T) system is clearly more accurate than an atom-centered invariant point charge MM model, for example, because it includes electronic polarization; thus the effects of the "link" atoms at the QM/MM boundary will be minimized. Furthermore, projector embedding can be used in situations where subsystems cut across all types of bonding. In short, the projector-embedding technique as applied in combination with QM/MM calculations provides an attractive approach for calculations on large biomolecular systems, because it has a simple formulation, is flexible, and makes possible the routine calculation of CCSD(T)-level energies (e.g., energy profiles for reactions) and other electronic properties.

AUTHOR INFORMATION

Corresponding Authors

*E-mail (S.J.B.): sb13343@bristol.ac.uk.
*E-mail (F.R.M.): fred.manby@bristol.ac.uk.
*E-mail (A.J.M.): adrian.mulholland@bristol.ac.uk.

Present Address

#M.S.: Physics Department, King's College London, Strand, London WC2R 2LS UK.

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

Author Contributions

The work to couple QM/MM with projector embedding was initiated by S.J.B. and M.W.v.d.K., embedding calculations were performed by S.J.B., R.C.R.P. and M.S. The initial programming to couple the MM potential to projector embedding was done by S.J.B. and M.S. All authors contributed to the analysis of results and the writing of the manuscript and have given approval to the final version of the manuscript. Note: All data from this work are held in an open-access repository.[109]

Funding

We are grateful to EPSRC for funding for this work through research grants (Grant No. EP/K018965/1 and EP/J012742/1) and the Doctoral Training Grant to the University of Bristol. M.W.vdK. is a BBSRC David Phillips Fellow and (with A.J.M.) thanks BBSRC for support (BB/M026280/1 and BB/L018756/1).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708−2728.

(2) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601−3658.

(3) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(4) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chem., Int. Ed.* **2006**, *45*, 6856−6859.

(5) Daniels, A. D.; Campeotto, I.; van der Kamp, M. W.; Bolt, A. H.; Trinh, C. H.; Phillips, S. E. V; Pearson, A. R.; Nelson, A.; Mulholland, A. J.; Berry, A. Reaction Mechanism of N -Acetylneuraminic Acid Lyase Revealed by a Combination of Crystallography, QM/MM Simulation, and Mutagenesis. *ACS Chem. Biol.* **2014**, *9*, 1025−1032.

(6) Kazemi, M.; Himo, F.; Åqvist, J. Enzyme Catalysis by Entropy without Circe Effect. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2406−2411.

(7) van der Kamp, M. W.; Żurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Testing High-Level QM/MM Methods for Modeling Enzyme Reactions: Acetyl-CoA Deprotonation in Citrate Synthase. *J. Phys. Chem. B* **2010**, *114*, 11303−11314.

(8) Karton, A.; O'Reilly, R. J.; Radom, L. Assessment of Theoretical Procedures for Calculating Barrier Heights for a Diverse Set of Water-Catalyzed Proton-Transfer Reactions. *J. Phys. Chem. A* **2012**, *116*, 4211−4221.

(9) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. High-Level QM/MM Modelling Predicts an Arginine as the Acid in the Condensation Reaction Catalysed by Citrate Synthase. *Chem. Commun.* **2008**, *No. 16*, 1874−1876.

(10) Oláh, J.; Harvey, J. N. NO Bonding to Heme Groups: DFT and Correlated Ab Initio Calculations. *J. Phys. Chem. A* **2009**, *113*, 7338−7345.

(11) Strickland, N.; Harvey, J. N. Spin-Forbidden Ligand Binding to the Ferrous−Heme Group: Ab Initio and DFT Studies. *J. Phys. Chem. B* **2007**, *111*, 841−852.

(12) Mata, R. A.; Werner, H.-J.; Thiel, S.; Thiel, W. Toward Accurate Barriers for Enzymatic Reactions: QM/MM Case Study on P-Hydroxybenzoate Hydroxylase. *J. Chem. Phys.* **2008**, *128*, 025104.

(13) Grimme, S. Improved Second-Order Møller−Plesset Perturbation Theory by Separate Scaling of Parallel- and Antiparallel-Spin Pair Correlation Energies. *J. Chem. Phys.* **2003**, *118*, 9095−9102.

(14) Mlýnský, V.; Banáš, P.; Šponer, J.; van der Kamp, M. W.; Mulholland, A. J.; Otyepka, M. Comparison of Ab Initio, DFT, and Semiempirical QM/MM Approaches for Description of Catalytic Mechanism of Hairpin Ribozyme. *J. Chem. Theory Comput.* **2014**, *10*, 1608−1622.

(15) Kaiyawet, N.; Lonsdale, R.; Rungrotmongkol, T.; Mulholland, A. J.; Hannongbua, S. High-Level QM/MM Calculations Support the Concerted Mechanism for Michael Addition and Covalent Complex Formation in Thymidylate Synthase. *J. Chem. Theory Comput.* **2015**, *11*, 713−722.

(16) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange−correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51−57.

(17) Friesner, R. A.; Guallar, V. Ab Initio Quantum Chemical and Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389−427.

(18) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(19) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(20) Hafner, J. Ab-Initio Simulations of Materials Using VASP: Density-Functional Theory and beyond. *J. Comput. Chem.* **2008**, *29*, 2044−2078.

(21) Neese, F. Prediction of Molecular Properties and Molecular Spectroscopy with Density Functional Theory: From Fundamental Theory to Exchange-Coupling. *Coord. Chem. Rev.* **2009**, *253*, 526−563.

(22) Pribram-Jones, A.; Gross, D. A.; Burke, K. DFT: A Theory Full of Holes? *Annu. Rev. Phys. Chem.* **2015**, *66*, 283−304.

(23) Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757−10816.

(24) Jakobsen, S.; Kristensen, K.; Jensen, F. Electrostatic Potential of Insulin: Exploring the Limitations of Density Functional Theory and Force Field Methods. *J. Chem. Theory Comput.* **2013**, *9*, 3978−3985.

(25) Lawan, N.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J. Comparison of DFT and Ab Initio QM/MM Methods for Modelling Reaction in Chorismate Synthase. *Chem. Phys. Lett.* **2014**, *608*, 380−385.

(26) Neese, F. Prediction of Electron Paramagnetic Resonance G Values Using Coupled Perturbed Hartree−Fock and Kohn−Sham Theory. *J. Chem. Phys.* **2001**, *115*, 11080−11096.

(27) Kaupp, M.; Reviakine, R.; Malkina, O. L.; Arbuznikov, A.; Schimmelpfennig, B.; Malkin, V. G. Calculation of Electronic G-Tensors for Transition Metal Complexes Using Hybrid Density Functionals and Atomic Meanfield Spin-Orbit Operators. *J. Comput. Chem.* **2002**, *23*, 794−803.

(28) Bennie, S. J.; Collison, D.; McDouall, J. J. W. Electronic and Magnetic Properties of Kremer's Tris-Hydroxo Bridged Chromium Dimer: A Challenge for DFT. *J. Chem. Theory Comput.* **2012**, *8*, 4915−4921.

(29) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.

(30) Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211−228.

(31) Mardirossian, N.; Head-Gordon, M. Exploring the Limit of Accuracy for Density Functionals Based on the Generalized Gradient Approximation: Local, Global Hybrid, and Range-Separated Hybrid Functionals with and without Dispersion Corrections. *J. Chem. Phys.* **2014**, *140*, 18A527.

(32) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. Effects of Dispersion in Density Functional Based Quantum Mechanical/Molecular Mechanical Calculations on Cytochrome P450 Catalyzed Reactions. *J. Chem. Theory Comput.* **2012**, *8*, 4637−4645.

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

(33) Rydberg, P.; Lonsdale, R.; Harvey, J. N.; Mulholland, A. J.; Olsen, L. Trends in Predicted Chemoselectivity of Cytochrome P450 Oxidation: B3LYP Barrier Heights for Epoxidation and Hydroxylation Reactions. *J. Mol. Graphics Modell.* **2014**, *52*, 30−35.
(34) Zhang, H.-M.; Chen, S.-L. Include Dispersion in Quantum Chemical Modeling of Enzymatic Reactions: The Case of Isoaspartyl Dipeptidase. *J. Chem. Theory Comput.* **2015**, *11*, 2525−2535.
(35) Bally, T.; Sastry, G. N. Incorrect Dissociation Behavior of Radical Ions in Density Functional Calculations. *J. Phys. Chem. A* **1997**, *101*, 7923−7925.
(36) Lundberg, M.; Siegbahn, P. E. M. Quantifying the Effects of the Self-Interaction Error in DFT: When Do the Delocalized States Appear? *J. Chem. Phys.* **2005**, *122*, 224103.
(37) Kim, M.-C.; Sim, E.; Burke, K. Ions in Solution: Density Corrected Density Functional Theory (DC-DFT). *J. Chem. Phys.* **2014**, *140*, 18A528.
(38) Mintmire, J. W.; Dunlap, B. I. Fitting the Coulomb Potential Variationally in Linear-Combination-of-Atomic-Orbitals Density-Functional Calculations. *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *25*, 88−95.
(39) Dunlap, B. I. Robust and Variational Fitting: Removing the Four-Center Integrals from Center Stage in Quantum Chemistry. *J. Mol. Struct.: THEOCHEM* **2000**, *529*, 37−40.
(40) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP: Linear-Scaling Density Functional Simulations on Parallel Computers. *J. Chem. Phys.* **2005**, *122*, 084119.
(41) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.
(42) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(43) Senthilkumar, K.; Mujika, J. I.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J.; Harvey, J. N. Analysis of Polarization in QM/MM Modelling of Biologically Relevant Hydrogen Bonds. *J. R. Soc., Interface* **2008**, *5*, 207−216.
(44) Pentikäinen, U.; Shaw, K. E.; Senthilkumar, K.; Woods, C. J.; Mulholland, A. J. Lennard−Jones Parameters for B3LYP/CHARMM27 QM/MM Modeling of Nucleic Acid Bases. *J. Chem. Theory Comput.* **2009**, *5*, 396−410.
(45) Shaw, K. E.; Woods, C. J.; Mulholland, A. J. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. *J. Phys. Chem. Lett.* **2010**, *1*, 219−223.
(46) Boulanger, E.; Thiel, W. Toward QM/MM Simulation of Enzymatic Reactions with the Drude Oscillator Polarizable Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 1795−1809.
(47) Steinmann, C.; Olsen, J. M. H.; Kongsted, J. Nuclear Magnetic Shielding Constants from Quantum Mechanical/Molecular Mechanical Calculations Using Polarizable Embedding: Role of the Embedding Potential. *J. Chem. Theory Comput.* **2014**, *10*, 981−988.
(48) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344−5355.
(49) Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761−777.
(50) Lever, G.; Cole, D. J.; Lonsdale, R.; Ranaghan, K. E.; Wales, D. J.; Mulholland, A. J.; Skylaris, C.-K.; Payne, M. C. Large-Scale Density Functional Theory Transition State Searching in Enzymes. *J. Phys. Chem. Lett.* **2014**, *5*, 3614−3619.
(51) Hampel, C.; Werner, H.-J. Local Treatment of Electron Correlation in Coupled Cluster Theory. *J. Chem. Phys.* **1996**, *104*, 6286−6297.
(52) Yang, J.; Chan, G. K.-L.; Manby, F. R.; Schütz, M.; Werner, H.-J. The Orbital-Specific-Virtual Local Coupled Cluster Singles and Doubles Method. *J. Chem. Phys.* **2012**, *136*, 144105.
(53) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and Accurate Approximations to the Local Coupled Cluster Singles Doubles Method Using a Truncated Pair Natural Orbital Basis. *J. Chem. Phys.* **2009**, *131*, 064103.
(54) Lykos, P. G.; Parr, R. G. On the Pi-Electron Approximation and Its Possible Refinement. *J. Chem. Phys.* **1956**, *24*, 1166−1173.
(55) Knizia, G.; Chan, G. K.-L. Density Matrix Embedding: A Strong-Coupling Quantum Embedding Theory. *J. Chem. Theory Comput.* **2013**, *9*, 1428−1432.
(56) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; Miller, T. F., III Embedded Mean-Field Theory. *J. Chem. Theory Comput.* **2015**, *11*, 568−580.
(57) Zgid, D.; Chan, G. K.-L. Dynamical Mean-Field Theory from a Quantum Chemical Perspective. *J. Chem. Phys.* **2011**, *134*, 094115.
(58) Wesolowski, T. A.; Warshel, A. Frozen Density Functional Approach for Ab Initio Calculations of Solvated Molecules. *J. Phys. Chem.* **1993**, *97*, 8050−8053.
(59) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III Exact Nonadditive Kinetic Potentials for Embedded Density Functional Theory. *J. Chem. Phys.* **2010**, *133*, 084103.
(60) Goodpaster, J. D.; Barnes, T. A.; Miller, T. F., III Embedded Density Functional Theory for Covalently Bonded and Strongly Interacting Subsystems. *J. Chem. Phys.* **2011**, *134*, 164108.
(61) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate Frozen-Density Embedding Potentials as a First Step towards a Subsystem Description of Covalent Bonds. *J. Chem. Phys.* **2010**, *132*, 164101.
(62) Nafziger, J.; Wu, Q.; Wasserman, A. Molecular Binding Energies from Partition Density Functional Theory. *J. Chem. Phys.* **2011**, *135*, 234101.
(63) Huang, C.; Pavone, M.; Carter, E. A. Quantum Mechanical Embedding Theory Based on a Unique Embedding Potential. *J. Chem. Phys.* **2011**, *134*, 154110.
(64) Mata, R. A.; Werner, H. J.; Schütz, M. Correlation Regions within a Localized Molecular Orbital Approach. *J. Chem. Phys.* **2008**, *128*, 144106.
(65) Mulholland, A. J.; Lyne, P. D.; Karplus, M. Ab Initio QM/MM Study of the Citrate Synthase Mechanism. A Low-Barrier Hydrogen Bond Is Not Involved. *J. Am. Chem. Soc.* **2000**, *122*, 534−535.
(66) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. Substrate Polarization in Enzyme Catalysis: QM/MM Analysis of the Effect of Oxaloacetate Polarization on Acetyl-CoA Enolization in Citrate Synthase. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 521−535.
(67) van der Kamp, M. W.; McGeagh, J. D.; Mulholland, A. J. Lethal Synthesis" of Fluorocitrate by Citrate Synthase Explained through QM/MM Modeling. *Angew. Chem., Int. Ed.* **2011**, *50*, 10349−10351.
(68) Karpusas, M.; Holland, D.; Remington, S. J. 1.9-Å Structures of Ternary Complexes of Citrate Synthase with D- and L-Malate: Mechanistic Implications. *Biochemistry* **1991**, *30*, 6024−6031.
(69) Remington, S. J. Structure and Mechanism of Citrate Synthase. *Curr. Top. Cell. Regul.* **1992**, *33*, 209−229.
(70) Evans, C. T.; Kurz, L. C.; Remington, S. J.; Srere, P. A. Active Site Mutants of Pig Citrate Synthase: Effects of Mutations on the Enzyme Catalytic and Structural Properties. *Biochemistry* **1996**, *35*, 10661−10672.
(71) Alter, G. M.; Casazza, J. P.; Zhi, W.; Nemeth, P.; Srere, P. A.; Evans, C. T. Mutation of Essential Catalytic Residues in Pig Citrate Synthase. *Biochemistry* **1990**, *29*, 7557−7563.
(72) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F., III A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564−2568.
(73) Wang, Y. A.; Govind, N.; Carter, E., A. Orbital-free kinetic-energy density functionals with a density-dependent kernel. *Phys. Rev. B* **1999**, *60*, 16350.
(74) Peach, M. J. G.; Griffiths, D. G. J.; Tozer, D. J. On the Evaluation of the Non-Interacting Kinetic Energy in Density Functional Theory. *J. Chem. Phys.* **2012**, *136*, 144101.
(75) Borgoo, A.; Tozer, D. J. Density Scaling of Noninteracting Kinetic Energy Functionals. *J. Chem. Theory Comput.* **2013**, *9*, 2250−2255.
(76) Borgoo, A.; Teale, A. M.; Tozer, D. J. Revisiting the Density Scaling of the Non-Interacting Kinetic Energy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14578−14583.

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689−2697

(77) Borgoo, A.; Green, J. A.; Tozer, D. J. Molecular Binding in Post-Kohn–Sham Orbital-Free DFT. *J. Chem. Theory Comput.* **2014**, *10*, 5338–5345.
(78) Bennie, S. J.; Stella, M.; Miller, T. F., III; Manby, F. R. Accelerating Wavefunction in Density-Functional-Theory Embedding by Truncating the Active Basis Set. *J. Chem. Phys.* **2015**, *143*, 024105.
(79) Stella, M.; Bennie, S. J.; Manby, F. R. Computational Study of Adsorption of Cobalt on Benzene and Coronene. *Mol. Phys.* **2015**, *113*, 1858–1864.
(80) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F., III Accurate and Systematically Improvable Density Functional Theory Embedding for Correlated Wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.
(81) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F., III Accurate Basis Set Truncation for Wavefunction Embedding. *J. Chem. Phys.* **2013**, *139*, 024103.
(82) Barnes, T. A.; Kaminski, J. W.; Borodin, O.; Miller, T. F., III Ab Initio Characterization of the Electrochemical Stability and Solvation Properties of Condensed-Phase Ethylene Carbonate and Dimethyl Carbonate Mixtures. *J. Phys. Chem. C* **2015**, *119*, 3865–3880.
(83) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
(84) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Györffy, W.; Kats, D.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M. *MOLPRO, a Package of ab Initio Programs*, version 2015.1; TTI GmbH, 2015; http://www.molpro.net.
(85) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 242–253.
(86) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
(87) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
(88) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
(89) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.
(90) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
(91) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(92) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(93) Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
(94) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
(95) Purvis, G. D.; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
(96) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
(97) Varetto, U. *Molekel 5.4* http://ugovaretto.github.io/molekel/ (accessed Jan 15, 2016).
(98) Schrödinger, L. *The PyMOL Molecular Graphics System*, version 1.8; Schrödinger, LLC, 2015.
(99) Werner, H.-J.; Manby, F. R. Explicitly Correlated Second-Order Perturbation Theory Using Density Fitting and Local Approximations. *J. Chem. Phys.* **2006**, *124*, 054114.
(100) Werner, H.-J. Eliminating the Domain Error in Local Explicitly Correlated Second-Order Møller–Plesset Perturbation Theory. *J. Chem. Phys.* **2008**, *129*, 101103.
(101) Adler, T. B.; Werner, H.-J.; Manby, F. R. Local Explicitly Correlated Second-Order Perturbation Theory for the Accurate Treatment of Large Molecules. *J. Chem. Phys.* **2009**, *130*, 054106.
(102) Adler, T. B.; Werner, H.-J. Local Explicitly Correlated Coupled-Cluster Methods: Efficient Removal of the Basis Set Incompleteness and Domain Errors. *J. Chem. Phys.* **2009**, *130*, 241101.
(103) Schütz, M.; Werner, H.-J. Local Perturbative Triples Correction (T) with Linear Cost Scaling. *Chem. Phys. Lett.* **2000**, *318*, 370–378.
(104) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-Order Scaling Local Electron Correlation Methods. I. Linear Scaling Local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.
(105) Schütz, M.; Manby, F. R. Linear Scaling Local Coupled Cluster Theory with Density Fitting. Part I: 4-External Integrals. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3349–3358.
(106) Werner, H.-J.; Schütz, M. An Efficient Local Coupled Cluster Method for Accurate Thermochemistry of Large Systems. *J. Chem. Phys.* **2011**, *135*, 144116.
(107) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.
(108) Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007**, *127*, 221106.
(109) Open access data held in the repository at https://dx.doi.org/10.5523/bris.2elwx7gwh5of1cean5iht89cm.

DOI: 10.1021/acs.jctc.6b00285
J. Chem. Theory Comput. 2016, 12, 2689–2697